LEARNING HIGHLIGHTS USING EVENT DETECTION

CROSS-REFERENCE TO RELATED APPLICATIONS

The application is a continuation of U.S. patent application Ser. No. 14/585,075, filed Dec. 29, 2014, which is a continuation of U.S. patent application Ser. No. 13/314,837, filed Dec. 8, 2011, which claims the benefit of U.S. Provisional Application No. 61/421,145, filed Dec. 8, 2010, each of which is hereby incorporated by reference herein in its entirety.

TECHNICAL FIELD

The present disclosure generally relates to automatically identifying video highlights.

BACKGROUND

Sports videos are punctuated by moments of excitement. For many sports, the exciting moments are scattered throughout a full video of the game which typically includes primarily uninteresting material. For every home run there are balls and strikes. For every touchdown or interception, unproductive running plays and incomplete passes. Soccer and hockey can have entire games with just a few goals. Most viewers just want to get these interesting and exciting portions (herein "sports video highlights"), without having to watch an entire game.

SUMMARY

A highlight learning module trains highlight classifiers to identify highlights in videos based on event vectors which characterize the videos according to detected events. To identify the events, features are extracted from the videos on a frame basis. The features are used to identify events within the video using event models trained in an unsupervised manner to identify recurring events within the videos. Using the event framework, the training videos are transcribed into a series of events and event vectors are constructed for the training videos to train a classifier according to the event vectors. Since the event framework is developed with an unsupervised assessment of the low-level features, the only supervision which need be used in this technique is to designate the video at a high level for the training sets as highlight or non-highlight. Moreover, the low-level feature and event detection framework enables a system applicable to a wide variety of sports videos.

The highlight learning module is used to classify video clips using the trained classifiers. The highlight learning module receives a video, or portion of a video to be classified. The highlight learning module extracts features from the video or portion thereof to match the features used to train the event models. The extracted features from the video clips are transcribed by the event models. An event vector is created for the transcribed events, and the video is classified using the event vector applied to the highlight classifier to determine if the video is a highlight according to this highlight classifier. The same event vector for a video can be classified using several highlight classifiers which can determine whether the video belongs to any of the highlight types.

The features and advantages described in the specification are not all inclusive and, in particular, many additional features and advantages will be apparent to one of ordinary skill in the art in view of the drawings, specification, and claims. Moreover, it should be noted that the language used in the specification has been principally selected for readability and instructional purposes, and may not have been selected to delineate or circumscribe the inventive subject matter.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a block diagram of a video hosting service in which highlight learning can be employed according to an embodiment.

FIG. **2** illustrates the various components of a highlight learning module used in the video hosting service according to an embodiment.

FIG. **3** is a detailed view of the event modeling components according to an embodiment.

FIG. **4** is a data flow diagram showing iterative refinement of the event models.

FIG. **5** presents an overview of highlight detection using event modeling according to an embodiment.

The figures depict embodiments of the present disclosure for purposes of illustration only. One skilled in the art will readily recognize from the following description that alternative embodiments of the structures and methods illustrated herein may be employed without departing from the principles of the disclosure described herein.

DETAILED DESCRIPTION

System Architecture

FIG. **1** is a block diagram of a video hosting service **100** in which highlight learning with event modeling can be employed, according to one embodiment. The video hosting service **100** represents a system such as that of YOUTUBE that stores and provides videos to clients such as the client device **135**. The video hosting site **100** communicates with a plurality of content providers **130** and client devices **135** via a network **140** to facilitate sharing of video content between users. Note that in FIG. **1**, for the sake of clarity only one instance of content provider **130** and client **135** is shown, though there could be any number of each. The video hosting service **100** additionally includes a front end interface **102**, a video serving module **104**, a video search module **106**, an upload server **108**, a user database **114**, and a video repository **116**. Other conventional features, such as firewalls, load balancers, authentication servers, application servers, failover servers, site management tools, and so forth are not shown so as to more clearly illustrate the features of the video hosting site **100**. One example of a suitable site **100** is the YOUTUBE website, found at www.youtube.com. Other video hosting sites can be adapted to operate according to the teachings disclosed herein. The illustrated components of the video hosting website **100** can be implemented as single or multiple components of software or hardware. In general, functions described in one embodiment as being performed by one component can also be performed by other components in other embodiments, or by a combination of components. Furthermore, functions described in one embodiment as being performed by components of the video hosting website **100** can also be performed by one or more client devices **135** in other embodiments if appropriate.

Client devices **135** are computing devices that execute client software, e.g., a web browser or built-in client application, to connect to the front end interface **102** of the video